Appendix 17

Protein Science (1995), *4*:2261–2268. Cambridge University Press. Printed in the USA.

Structure comparison of native and mutant human recombinant FKBP12 complexes with the immunosuppressant drug FK506 (tacrolimus)

SUSUMU ITOH[1] AND MANUEL A. NAVIA

Vertex Pharmaceuticals Incorporated, Cambridge, Massachusetts 02139-4211

(RECEIVED June 13, 1995; ACCEPTED August 9, 1995)

Abstract

The consequences of site-directed mutagenesis experiments are often anticipated by empirical rules regarding the expected effects of a given amino acid substitution. Here, we examine the effects of "conservative" and "nonconservative" substitutions on the X-ray crystal structures of human recombinant FKBP12 mutants in complex with the immunosuppressant drug FK506 (tacrolimus). R42K and R42I mutant complexes show 110-fold and 180-fold decreased calcineurin (CN) inhibition, respectively, versus the native complex, yet retain full peptidyl prolyl isomerase (PPIase) activity, FK506 binding, and FK506-mediated PPIase inhibition. Interestingly, the structure of the R42I mutant complex is better conserved than that of the R42K mutant complex when compared to the native complex structure, within both the FKBP12 protein and FK506 ligand regions of the complexes, and with respect to temperature factors and RMS coordinate differences. This is due to compensatory interactions mediated by two newly ordered water molecules in the R42I complex structure, molecules that act as surrogates for the missing arginine guanidino nitrogens of R42. The absence of such surrogate solvent interactions in the R42K complex leads to some disorder in the so-called "40s loop" that encompasses the substituent. One rationalization proposed for the observed loss in CN inhibition in these R42 mutant complexes invokes indirect effects leading to a misorientation of FKBP12 and FK506 structural elements that normally interact with calcineurin. Our results with the structure of the R42I complex in particular suggest that the observed loss of CN inhibition might also be explained by the loss of a specific R42-mediated interaction with CN that cannot be mimicked effectively by the solvent molecules that otherwise stabilize the conformation of the 40s loop in that structure.

Keywords: calcineurin; immunophilins; site-directed mutagenesis; structure-based drug design; X-ray crystallography

FK506 (United States Adopted Names Council of the American Medical Association, Chicago, Illinois [USAN], *tacrolimus*) is a natural product screening lead (Kino et al., 1987) now approved for therapeutic use as an immunosuppressant in Japan, the USA, Germany, and other countries. FK506, in complex with its 12-kDa M_r binding protein FKBP12, exerts its immunosuppressive effects through the inhibition of calcineurin (CN), an intracellular Ca^{+2}-calmodulin-dependent phosphatase (Klee & Cohen, 1988). CN inhibition, in turn, interrupts the induction of IL-2 and other T-cell activation events (Friedman & Weissman, 1991; Liu et al., 1991). A homologous natural product, rapamycin (AY-22,989; USAN, *sirolimus*), which was initially discovered as an antifungal agent (Sehgal et al., 1975), can antagonize the CN inhibitory activity of FK506 (Bierer et al., 1990a; Dumont et al., 1990b), even though it is itself an immunosuppressant by a different mechanism (Bierer et al., 1990a; Dumont et al., 1990a). These differences in CN inhibitory activity between the agonist FK506 and the antagonist rapamycin in their complexes with FKBP12 (Table 1), were first explained within the framework of an elegant model (Schreiber, 1991) that focused attention on the corresponding differences in chemical structure between the two ligands in their so-called "effector domain" region (see, e.g., Fig. 1). X-ray crystallographic studies have provided support for this model (van Duyne et al., 1991a, 1991b, 1993; Becker et al., 1993; Rotonda et al., 1993; Connelly et al., 1994; Wilson et al., 1995), by showing that the effector

Reprint requests to: Manuel A. Navia, Vertex Pharmaceuticals Incorporated, 40 Allston Street, Cambridge, Massachusetts 02139-4211; e-mail: navia@vpharm.com.

[1] Visiting from: Chugai Pharmaceuticals Co. Ltd., 1-135 Komakado, Gotemba Shizuoka 412, Japan.

Abbreviations: CN, calcineurin; PPIase, peptidyl prolyl isomerase; FK506, USAN tacrolimus; FKBP12, 12-kDa M_r FK506 binding protein; FKBP13, 13-kDa M_r FK506 binding protein; USAN, United States Adopted Names Council of the American Medical Association, Chicago, Illinois.

Table 1. *Biochemical properties of the native and mutant complexes*[a]

Mutant	Calcineurin K_i (nM)	Reduction vs. native complex	FK506 K_i (nM)	Specific PPIase activity ($s^{-1}\ \mu M^{-1}$)
Native[b]	5.5 (1.8)	–	0.6 (0.2)	4.3 (0.4)
R42K[b]	590 (200)	107.3	0.6 (0.2)	3.8 (0.3)
R42I[b]	970 (150)	176.4	0.1 (0.1)	2.5 (0.3)
R42Q[c]	325 (150)	59.1	4.3 (2.0)	3.0 (0.3)
Native[d]	7.9 (3.0)	–	0.4 (0.2)	2.2 (0.2)
R42Q[d]	850 (250)	107.6	1.7 (0.6)	1.3 (0.3)
R42A[d]	280 (80)	35.4	0.2 (0.1)	1.1 (0.2)
Chimera[d,e]	19 (2)	2.4	0.4 (0.2)	0.57 (0.05)

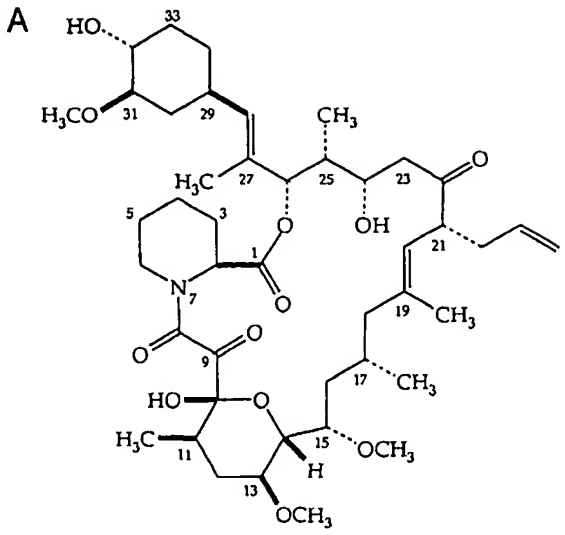
[a] Summary of published biochemical data for native and mutant FKBP12 proteins and their complexes with FK506 and calcineurin. The inhibition constant for calcineurin by native and mutant FKBP12 complexes with FK506 is given, along with the FKBP12 inhibition constants versus FK506 and the PPIase specific activity of the various FKBP12s versus a synthetic substrate.

[b] Aldape et al. (1992).

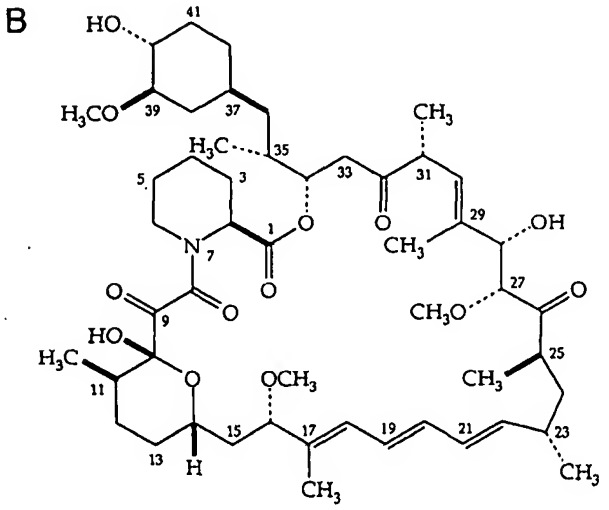
[c] Futer et al. (1995).

[d] Yang et al. (1993).

[e] Substitute FKBP12 residues 40–44 (-RDRNK-) with the corresponding residues (-LPQNQ-) from FKBP13.



FK-506



Rapamycin

Fig. 1. A: Chemical structure of FK506 (USAN, tacrolimus). The effector domain of FK506 (Schreiber, 1991) corresponds to those portions of the ligand (C18–C23 and substituents in the macrocycle, and C26–C34 in the cyclohexyl ring) that have been shown crystallographically to protrude from the surface of its complex with FKBP12; structural elements in common between rapamycin and FK506 have been shown crystallographically to bind the PPIase active site of FKBP12 in the same manner (van Duyne et al., 1991a, 1991b, 1993; Becker et al., 1993; Rotonda et al., 1993; Armistead et al., 1995; Itoh et al, 1995; Wilson et al., 1995). B: Chemical structure of rapamycin (AY-22989, USAN, sirolimus). The rapamycin effector domain corresponds to atoms C15–C29 in the macrocycle and C36–C42 in the cyclohexyl ring.

domains of FK506 and rapamycin do indeed protrude from the surface of their respective protein–ligand complexes (Fig. 2A) with distinct conformations that might be compatible with CN binding and inhibition for the one ligand, but not for the other. In turn, those chemical structure elements shared by the two ligands (Fig. 1) were shown in those studies to constitute FKBP12 "binding domains," allowing a rationalization of the reciprocal antagonism between the two ligands in terms of their competition for a common FKBP12 binding site.

The effector domain model has retained broad acceptance as a first approximation to the complicated problem of immunosuppressive drug design, in part because of its consistency with the observed loss of CN inhibitory activity that follows even minor variation in the chemical structure of FK506 (Goulet et al., 1994). By limiting the role of the FKBP12 protein to that of a presenter of ligand functionality to CN (Schreiber, 1991; Rosen & Schreiber, 1992; Schreiber et al., 1993), the model reduces the scope of the drug design problem to one of simple mimicry of the conformation of the FK506 effector domain that protrudes from the surface of the native FKBP12–FK506 complex. Unfortunately, these efforts have yet to produce a linear or macrocyclic drug lead—let alone a clinical candidate—that exceeds the potency of FK506 (Itoh et al., 1995); ligands predicted on the basis of this model have all turned out to be antagonists of FK506 (see, e.g., Bierer et al., 1990b; Somers et al., 1991; Armistead et al., 1995).

Experimental evidence for a more complicated interaction between CN and the FKBP12–FK506 complex first emerged from a systematic examination of the biochemical properties of site-directed mutants of charged residues on the surface of FKBP12 (Aldape et al., 1992). The critical involvement of the "40s loop" and "80s loop" regions of the protein (as defined in Fig. 2B) was established for this interaction by these and subsequent studies (Yang et al., 1993; Futer et al., 1995), leading to a generalization of the effector domain model in the direction of a composite "effector surface" of both protein and ligand structural elements. Elsewhere, we have explored the structural consequences that follow substitutions in the 80s loop region of FKBP12 (Itoh et al., 1995) and have identified composite features on the FKBP12–

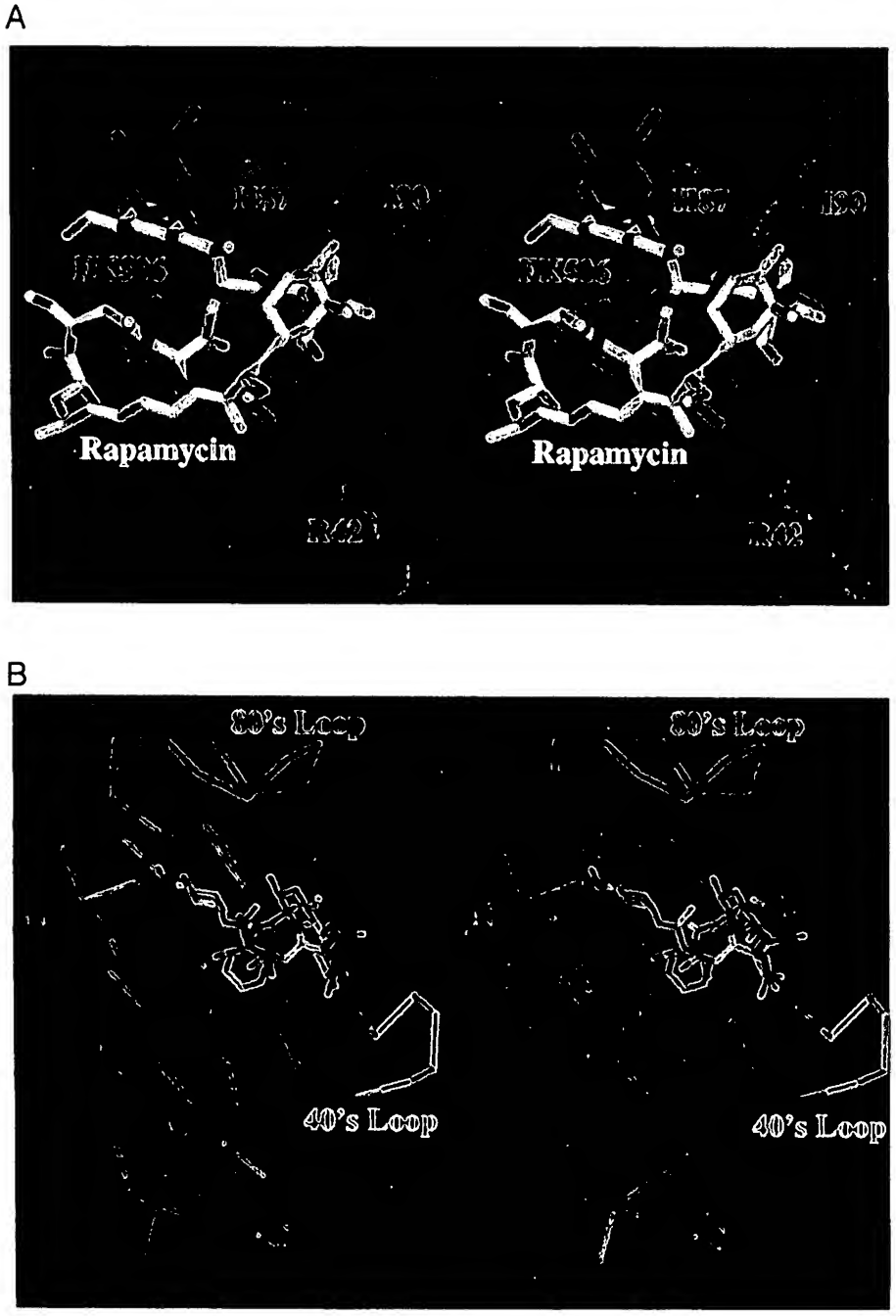


Fig. 2. **A:** Comparison of the structures of the FKBP12 complexes with FK506 (in red) and with rapamycin (in white). A dot representation of the surface of native FKBP12 (Wilson et al., 1995) is shown in blue and is representative of the surface of the FKBP12 complex structures with FK506 and rapamycin. **B:** Conformation of the backbone Cα of FKBP12 in its complex with FK506 is shown in blue. The 40s loop and 80s loop regions of the FKBP12 protein are specifically identified in yellow and red, respectively. The 40s loop is made up of residues 40–44 in FKBP12, which form a bulge on the third β-strand of the protein, as defined by van Duyne et al. (1991a). The 80s loop includes residues 84–91 on the edge of the fifth β-strand of the structure.

FK506 effector surface that might be candidate CN recognition and binding elements (Wilson et al., 1995). In this paper we focus on structural elements in the 40s loop region of FKBP12 and, in particular, on substitutions at residue 42 of the protein, which demonstrate profound (Table 1) but complicated effects on the CN inhibitory potency of the corresponding mutant complexes with FK506. These mutant data have led to proposed mechanisms of action, which are distinctly different in their character and consequences, that need to be resolved (Clardy, 1995).

Results

Crystallization, data collection, and refinement statistics for the native and mutant complex structures reported here are given in Table 2, along with RMS differences in conformation and mobility versus the native complex structure; biochemical data are summarized from the existing literature (Aldape et al., 1992; Yang et al., 1993; Futer et al., 1995) in Table 1. In all the complexes studied, the FKBP12 fold (Fig. 2B) that was seen in the

Table 2. *Summary of the wild-type and mutant complex structure analyses*[a]

	Wild type	R42K	R42I
Area detector used	Siemens	Rigaku	Rigaku
$P4_22_12$ unit cell; a, c (Å)	58.39, 55.76	58.31, 55.93	58.25, 55.98
Resolution (Å)	6.0–1.5	6.0–1.5	6.0–1.6
No. observations	76,344	43,803	44,188
% Reflection ($I > 2\sigma$)	81.4	88.1	89.7
R-merge (%)	3.99	5.88	3.15
R-factor (%)	16.6	18.7	16.8
No. water molecules	85	83	87
RMS bond length error (Å)	0.016	0.016	0.018
RMS bond angle error (deg)	2.74	2.85	2.93
Avg. FK506 B-factors (Å^2)	12.3	15.3	11.0
FK506 RMS diff vs. wt (Å)	–	0.139	0.123
FKBP12 RMS diff vs. wt (Å)	–	0.146	0.147

[a] Native and mutant FKBP12 complexes all share the native crystal form first reported by van Duyne et al. (1991a).

native complex structure (van Duyne et al., 1991a) is strongly conserved, consistent with the observed retention of PPIase and FK506 binding activity (Table 1). In this study, a considerable effort was made to crystallize all the complexes reported in a common crystal form (Table 2), in order to facilitate a direct comparison between structures. This crystal form turned out to be that of the native FKBP12-FK506 complex (van Duyne et al., 1991a), as a consequence of seeding mutant complex crystallization experiments with microcrystals of the native complex and subsequently using crystals from those solutions as macroseeds leading to data quality mutant complex crystals.

In the native complex structure (red coordinates in Fig. 3A, Kinemage 1), the two guanidino nitrogens of R42 are seen to stabilize the "40s loop" of FKBP12 through their participation in a bridging network of noncovalent interactions between residues D37 and K44. In the R42I mutant complex structure (yellow coordinates in Fig. 3A, Kinemage 1), two tightly bound water molecules (yellow in Fig. 3A, Kinemage 1) substitute for the missing

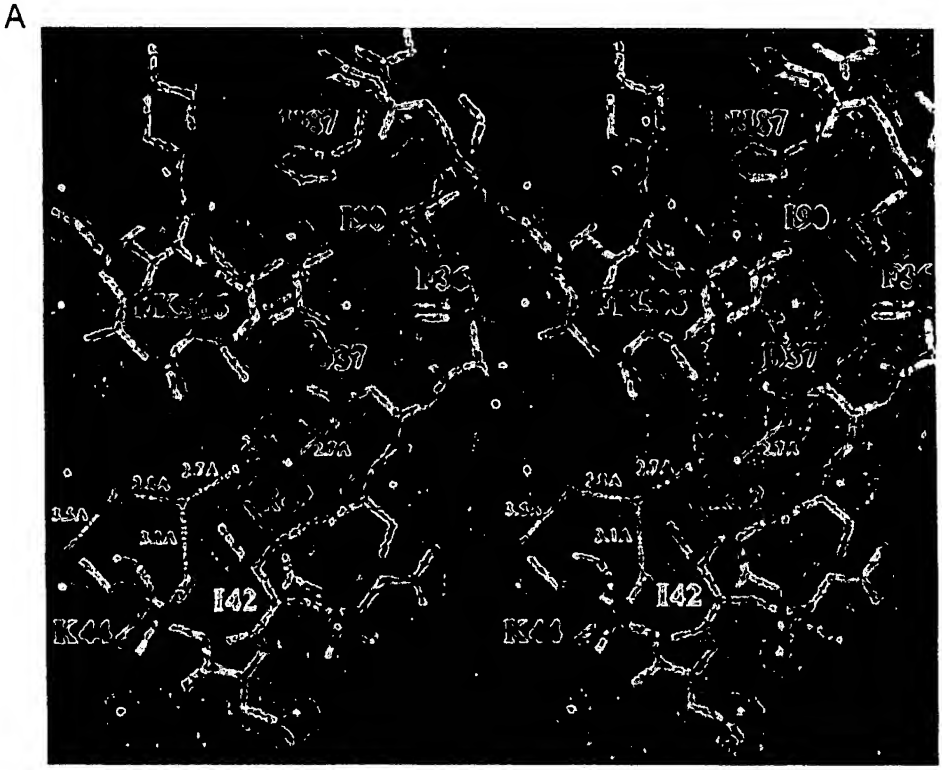
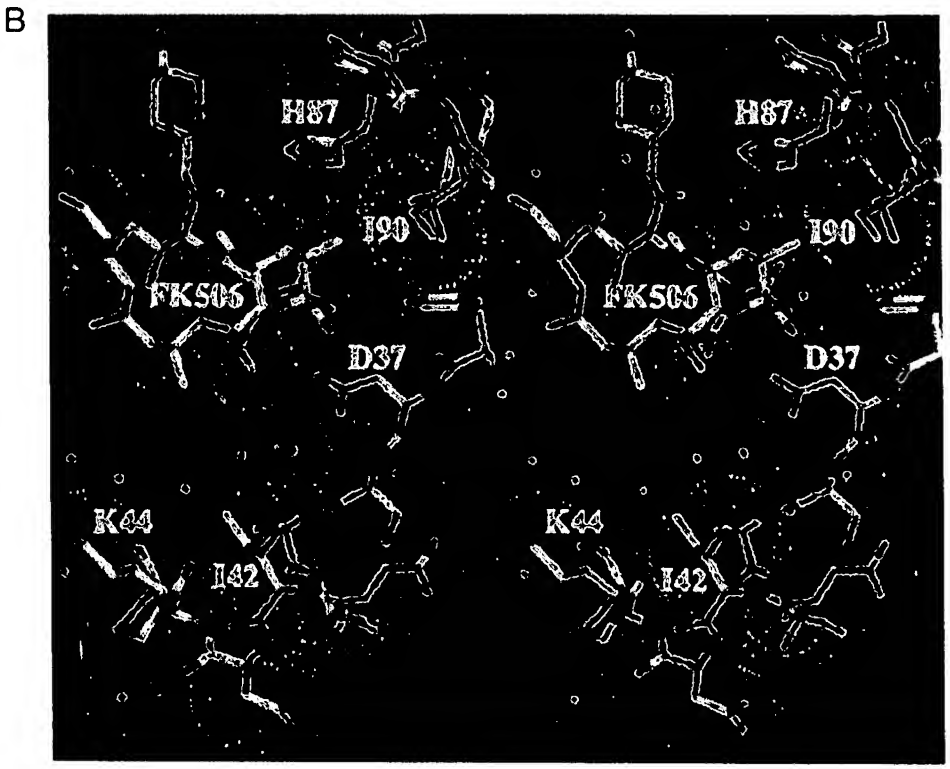


Fig. 3. **A:** Comparison of the structures of native (red) and R42I mutant (yellow) FKBP12 complexes with FK506. Refined coordinates (Table 2) of both complex structures are superimposed in the figure, together with the $2|F_o| - |F_c|$ electron density (in blue) of the R42I mutant complex structure, contoured at 1σ above background. Part of the FK506 binding domain is shown, as well as residues surrounding R42 and H87 in the protein. Water molecules bridging residues D37 and K44 are shown in green and yellow, corresponding to the R42I mutant complex structure. Dashed lines indicate bond distances between the water molecules. Our observations are inconsistent with the suggestion of Yang et al. (1993) that R42 mutants exert their effects indirectly, by reorienting nearby regions of the complex structure. Nor does the conformation of the 40s loop change as drastically as that of FKBP13 (Schultz et al., 1994) as a consequence of these substitutions. **B:** Dot-surface representation of the R42I mutant complex structure in the vicinity of FK506. Two water molecules (shown in green) in the R42I mutant complex structure fit readily into the gap created by the R42I substitution and act as surrogates for the missing guanidino nitrogens of R42 in bridging residues D37–K44. This interaction helps preserve the native conformation in the mutant complex and maintains PPIase and FK506 binding activity. Nonetheless, CN inhibition is drastically reduced (Table 1), suggesting a specific protein-protein interaction in the native complex that the water surrogates would be unable to mimic. **C:** Structure of the R42K mutant complex (in green) compared to that of the native complex (in red). As above, the $2|F_o| - |F_c|$ electron density corresponding to the R42K mutant complex, contoured at 1σ above background, is presented in blue. (*Continues on facing page.*)



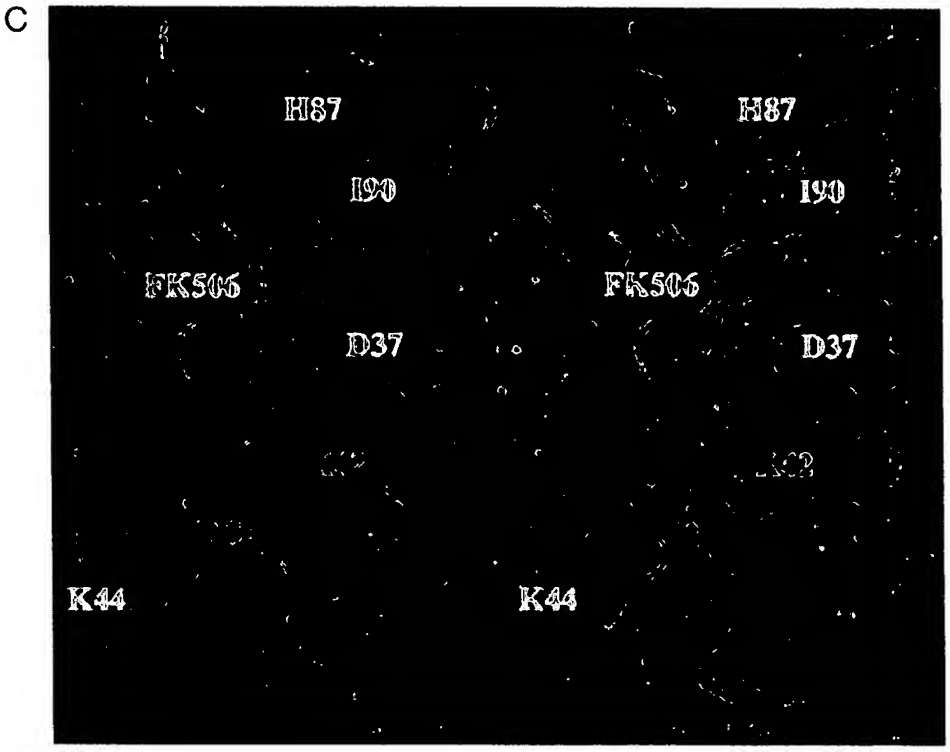


Fig. 3. *Continued.*

arginine guanidino nitrogens of R42, and are well accommodated (Fig. 3B, Kinemage 1) in the gap created by the smaller isoleucine substituent; a third water molecule (green in Fig. 3A, Kinemage 1) is apparently common to all the FKBP12 complex structures. As guanidino nitrogen surrogates, the two water molecules help to preserve the native complex 40s loop conformation in the R42I mutant complex, with an RMS deviation of 0.147 Å between the structures (Table 2). This interaction resembles one seen in the structure of a T157G mutant of T4 lysozyme (Alber et al., 1987; Matthews, 1993), where an ordered water molecule, acting as a surrogate for the missing T157 hydroxyl, preserves the pattern of stabilizing hydrogen bonding interactions seen in the native protein. It is interesting that, in spite of the high degree of conservation seen in the structure of the R42I mutant complex, CN inhibition is nonetheless reduced by ~180-fold (Table 1).

In the structure of the FK506 complex with the more conservative R42K mutant protein, the single ε-amino nitrogen of the lysine side chain is unable to substitute for both of the guanidino nitrogens of arginine (red in Fig. 3C, Kinemage 1); nor does the substitution leave enough space for an additional water molecule to insert itself as a surrogate, as shown in Figure 3B (Kinemage 1) for the R42I mutant complex. The resulting destabilization of the 40s loop is reflected in the higher temperature factors observed in that region of the protein (Fig. 4), even though the conformation of mutant protein complex still closely resembles that of the native (Fig. 3C, Kinemage 1), and FK506 binding and PPIase activity are preserved (Table 1).

Discussion

Two working models have emerged to rationalize the profound though complicated effects on CN inhibition that accompany substitutions in and around residue 42 of FKBP12. The simpler of these is made evident in the R42 single-site mutant complexes that were first characterized biochemically by Aldape et al. (1992), whose crystal structures are described here. In that model, the observed loss of CN-inhibitory activity can be immediately explained in terms of a direct and localized perturbation, by the substituted residue, of the effector surface presented to CN by the corresponding FKBP12–FK506 complex. The Merck group (Becker et al., 1993; Rotonda et al., 1993) arrives at a similar conclusion in their analysis and comparison of the structures of human and yeast FKBP12–FK506 complexes and of the human FKBP12 complex with L-685,818, an 18-hydroxy,21-ethyl analog of FK506.

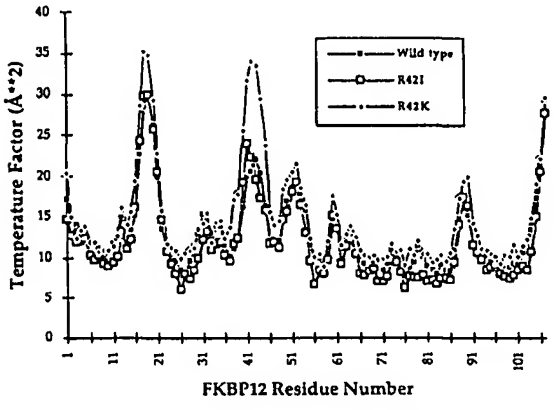


Fig. 4. Average X-ray temperature factors for the main-chain heavy atoms of FKBP12 for the native and the two mutant complex structures. Residues showing high temperature factors correspond to loop regions in the structure of FKBP12. These have been shown previously to be regions of local flexibility within the protein that were identified through a crystal structural analysis of 19 FKBP12–ligand complexes, each in a different crystal-packing arrangement (Wilson et al., 1995).

Yang et al. (1993), however, have suggested that substitutions at R42 exert their influence on CN inhibition indirectly, through a generalized conformational misorientation of nearby elements of the FKBP12–FK506 effector surface. This model is inferred from the curious pattern of CN-inhibitory activity evidenced in the FKBP12/13 chimeras studied by these workers. Substitution in FKBP12 of the corresponding 40s loop sequence from FKBP13 (i.e., replacing the sequence RDRNK with LPQNQ) leads to only a modest loss of CN-inhibitory potency (by ~2-fold to 19 nM). In turn, the single site R42Q mutant complex is severely compromised (by ~100-fold, to 850 nM), even though the R42Q substitution is incorporated in the FKBP12/13 chimera. From these observation, Yang et al. (1993) concluded that the effects of an R42 substitution would have to be strongly contextual. In other words, R42 and Q42 would each be appropriate to the 40s loop of FKBP12 and FKBP13 respectively, with only a modest loss of activity for the latter in the chimeric FKBP12/13 complex. An incompatible substitution, such as that of R42Q into FKBP12, would then lead to a significant and generalized disruption of the effector surface, an event that would be reflected in the much lower CN-inhibitory activity seen in the single-mutant complexes. Clardy (1995) has noted that the 40s loop in the structure of the native FKBP13–FK506 complex is displaced by about 2 Å RMS relative to the FKBP12–FK506 complex when these are overlapped, a point in support of the Yang et al. (1993) thesis.

The structural data presented here for the R42K and R42I mutant complexes show no such significant rearrangement of the 40s loop, the 80s loop, or any other part of the FKBP12 protein (Fig. 3). Nor do we observe a change in the conformation of FK506, even though we have demonstrated elsewhere (Itoh et al., 1995) that just such a conformational transformation is present in the FKBP12 R42K-H87V double mutant complex (Kinemage 2). All of our mutant complex structures (including the Itoh et al. [1995] double-mutant complex) have been solved in the same native FKBP12–FK506 complex crystal form described by van Duyne et al. (1991a), a crystal form that includes a significant number of ligand–ligand interactions (van Duyne et al., 1993; Wilson et al., 1995) that might otherwise have compromised the comparative interpretation we've presented for these structures. Even in the R42K mutant complex, where a significant increase in the temperature factors of the 40s loop of the mutant complex structure is observed (see Fig. 4), the weakened electron density in this region (Fig. 3C) is still consistent with the conformation of the 40s loop found in the native complex structure. In the less conservative R42I mutant complex structure, however, the fortuitous and unexpected ordering of two water molecules (Fig. 3A,B) acting as surrogates of the missing guanidino nitrogens of R42, leads to a more highly conserved structure, even though the loss of CN-inhibitory activity is also greater (~180-fold at 970 nM for the R42I mutant complex versus ~110-fold at 590 nM for the R42K mutant complex; Table 1).

Elsewhere, we describe the structures of FKBP12 (Itoh et al., 1995) and FKBP13 (Griffith JP, Wilson KP, Futer O, Living-

ston DJ, Navia MA. Structure of a mutant FKBP13-FK506 complex that is a high affinity inhibitor of calcineurin [manuscript in preparation]) mutant complexes that are inconsistent with the hypothesis that significant structural rearrangements in the 80s loop region of FKBP12 are responsible (Yang et al., 1993; Clardy, 1993, 1995) for the loss of CN-inhibitory activity seen in the corresponding FKBP12 mutant complexes. Nor do the results presented here support a similar hypothesis for the 40s loop region. Given our structural observations, one might speculate that R42 participates in some direct interaction with CN that cannot be effectively mimicked by the guanidino nitrogen surrogate water molecules that stabilize the 40s loop in the R42I mutant complex structure. With the R42K mutant complex, one might further consider an intermediate level of interaction with CN, through the single ϵ-amino nitrogen of the lysine substituent, and in spite of the greater disorder in the 40s loop. Teague (1995) has recently restated the importance of a conformationally well-defined recognition surface in promoting a strong interaction between exposed hydrophobic elements, such as are found in the FKBP12–FK506 effector surface and are presumed to exist on the complementary surface of CN. Our results show that relatively minor, localized perturbations of the CN complementarity of the FKBP12–FK506 effector surface can lead to quite significant effects on the CN-inhibitory potential of the resulting mutant complexes.

Methods

Mutant FKBP12 protein was prepared as reported (Aldape et al., 1992; Park et al., 1992; Wilson et al., 1995). Native and mutant FKBP12–FK506 complexes were prepared and crystals were grown essentially as described (van Duyne et al., 1991a; Wilson et al., 1995). Native crystals were used to seed the mutant complex crystallization experiments, and all the species reported here crystallized isomorphously in the native crystal form, as shown in Table 2. Diffraction data were collected on an X1000 multiwire area detector (Siemens Analytical Instruments, Madison, Wisconsin) or on an R-Axis II image plate detector (Rigaku/MSC, Woodlands, Texas), as indicated in Table 2. Data collection and processing used software provided by the manufacturers. All data were collected at room temperature. The reported structure of native FKBP12 in complex with FK506 (van Duyne et al., 1991a; Brookhaven Protein Data Bank [Bernstein et al., 1977] entry 1FKF) was used directly as an initial model for the crystallographic refinement of the mutant and native protein complexes. Refinement was by simulated annealing using the X-PLOR program package (Brünger, 1992). Mutated amino acids were initially refined as alanine, and the actual mutant side chains were introduced as refinement progressed. Water molecules were positioned in the model with the aid of a peak search program (SERC Daresbury Laboratory, 1979). The program QUANTA (Molecular Simulations; Burlington, Massachusetts) was used to examine electron density maps and protein models and for the superposition of structures and the calculation of the RMS differences reported (Table 2). Coordinates for the structures are being deposited in the Protein Data Bank.

Acknowledgments

We thank our colleagues Robert A. Aldape, Steven Chambers, Paul Caron, Maureen DeCenzo, John Fulghum, Olga Futer, James Griffith, David Livingston, and Mark Murcko for their continued assistance and support. In particular, Maureen DeCenzo and David Livingston provided the native and mutant proteins used in these studies. Paul Caron prepared the Kinemage files that accompany this paper. We also thank Joshua Boger, Patrick Connelly, Christopher Lepre, and Vicki Sato for their comments on this manuscript.

References

Alber T, Dao-pin S, Wilson K, Wozniak JA, Cook SP, Matthews BW. 1987. Contributions of hydrogen bonds of Thr 157 to the thermodynamic stability of phage T4 lysozyme. *Nature (Lond) 330*:41–46.

Aldape RA, Futer O, DeCenzo MT, Jarrett BP, Murcko MA, Livingston DJ. 1992. Charged surface residues of FKBP12 participate in formation of the FKBP12–FK506 calcineurin complex. *J Biol Chem 267*:16029–16032.

Armistead DA, Badia MC, Deininger DD, Duffy JP, Saunders, JO, Tung RD, Thomson JA, DeCenzo MT, Futer O, Livingston DJ, Murcko MA, Yamashita MM, Navia MA. 1995. Design, synthesis and structure of nonmacrocyclic inhibitors of FKBP12, the major binding protein for the immunosuppressant FK506. *Acta Crystallogr D51*:522–528.

Becker JW, Rotonda J, McKeever BM, Chan HK, Marcy AI, Wiederrecht G, Hermes JD, Springer JP. 1993. FK-506-binding protein: Three dimensional structure of the complex with the antagonist L-685,818. *J Biol Chem 268*:11335–11339.

Bernstein FC, Koetzel TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Bierer BE, Mattila PS, Standaert RF, Herzenberg LA, Burakoff SJ, Crabtree G, Schreiber SL. 1990a. Two distinct signal transmission pathways in T-lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin. *Proc Natl Acad Sci USA 87*: 9231–9235.

Bierer BE, Somers PK, Wandless TJ, Burakoff SJ, Schreiber SL. 1990b. Probing immunosuppressant action with a nonnatural immunophilin ligand. *Science 250*:556–559.

Brünger AT. 1992. *X-PLOR version 3.1: A system for X-ray crystallography and NMR*. New Haven, Connecticut: Yale University.

Clardy J. 1993. Structural studies of complexed FK-506 binding protein. *Ann NY Acad Sci 685*:37–46.

Clardy J. 1995. The chemistry of signal transduction. *Proc Natl Acad Sci USA 92*:56–61.

Connelly PR, Aldape RA, Bruzzese FJ, Chambers SP, Fitzgibbon MJ, Fleming MA, Itoh S, Livingston DJ, Navia MA, Thomson JA, Wilson KP. 1994. Enthalpy of hydrogen bond formation in a protein-ligand binding reaction. *Proc Natl Acad Sci USA 91*:1964–1968.

Dumont FJ, Melino MR, Staruch MJ, Koprak SL, Fischer PA, Sigal NH. 1990a. The immunosuppressive macrolides FK-506 and rapamycin act as reciprocal antagonists in murine T cells. *J Immunol 144*:1418–1424.

Dumont FJ, Staruch MJ, Koprak SL, Melino MR, Sigal NH. 1990b. Distinct mechanisms of suppression of murine T cell activation by the related macrolides FK-506 and rapamycin. *J Immunol 144*:251–258.

Friedman J, Weissman I. 1991. Two cytoplasmic candidates for immunophilin action are revealed by affinity for a new cyclophilin; one in the presence and one in the absence of CsA. *Cell 66*:799–806.

Futer O, DeCenzo MT, Aldape RA, Livingston DJ. 1995. FK506 binding protein mutational analysis: Defining the surface residue contributions to stability of the calcineurin co-complex. *J Biol Chem 270*:18935–18940.

Goulet MT, Rupprecht KM, Sinclair PJ, Wyvratt MJ, Parsons WH. 1994. The medicinal chemistry of FK-506. *Perspect Drug Disc Design 2*:145–162.

Itoh S, DeCenzo MT, Livingston DJ, Pearlman DA, Navia MA. 1995. Conformation of FK506 in X-ray structures of its complexes with human recombinant FKBP12 mutants. *Bioorg Med Chem Lett 5*:1983–1988.

Kino T, Hatanaka H, Hashimoto M, Nishiyama M, Goto T, Okuhara M, Kohsaka M, Aoki H, Imanaka H. 1987. FK-506, a novel immunosuppressant isolated from a *Streptomyces*. I. fermentation, isolation and physico-chemical and biological characteristics. *J Antibiot (Tokyo) 40*: 1249–1255.

Klee CB, Cohen P. 1988. The calmodulin-regulated protein phosphatase. *Mol Aspects Cell Regul 5*:225–248.

Liu J, Farmer JD Jr, Lane WS, Friedman J, Weissman I, Schreiber SL. 1991. Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP–FK506 complexes. *Cell 66*:807–815.

Matthews BW. 1993. Structural and genetic analysis of protein stability. *Annu Rev Biochem 62*:139–160.

Park ST, Aldape RA, Futer O, DeCenzo MT, Livingston DJ. 1992. PPIase

catalysis by human FK506-binding protein proceeds through a conformational twist mechanism. *J Biol Chem 267*:3316–3324.
Rosen MK, Schreiber SL. 1992. Natural products as probes of cellular function: Studies on immunophilins. *Angew Chem Intl Ed Engl 31*:384–400.
Rotonda J, Burbaum JJ, Chan KH, Marcy AI, Becker JW. 1993. Improved calcineurin inhibition by yeast FKBP12-drug complexes: Crystallographic and functional analysis. *J Biol Chem 268*:7607–7609.
Schreiber SL. 1991. Chemistry and biology of the immunophilins and their immunosuppressive ligands. *Science 251*:283–287.
Schreiber SL, Albers MW, Brown EJ. 1993. The cell cycle, signal transduction, and immunophilin-ligand complexes. *Acc Chem Res 26*:412–420.
Schultz LW, Martin PK, Liang J, Schreiber SL, Clardy J. 1994. Atomic structure of the immunophilin FKBP13-FK506 complex: Insights into the composite binding surface for calcineurin. *J Am Chem Soc 116*:3129–3130.
Sehgal SN, Baker H, Vezina C. 1975. Rapamycin (AY-22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomycete and isolation of the active principal. *J Antibiot (Tokyo) 28*:727–732.
SERC Daresbury Laboratory. 1979. *CCP4: Collaborative computing project no. 4. A suite of programs for protein crystallography.* Warrington WA4 4AD, UK: Daresbury Laboratory.
Somers PK, Wandless TJ, Schreiber SL. 1991. Synthesis and analysis of 506BD, a high-affinity ligand for the immunophilin FKBP. *J Am Chem Soc 113*:8045–8056.
Teague S. 1995. Lessons from molecular matchmakers. *Struct Biol 2*: 360–361.
van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J. 1991a. Atomic structure of FKBP-FK-506, an immunophilin-immunosuppressant complex. *Science 252*:839–842.
van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J. 1993. Atomic structures of human immunophilin FKBP12 complexes with FK506 and rapamycin. *J Mol Biol 229*:105–124.
van Duyne GD, Standaert RF, Schreiber SL, Clardy J. 1991b. Atomic structure of the rapamycin human immunophilin FKBP-12 complex. *J Am Chem Soc 113*:7433–7434.
Wilson KP, Yamashita MM, Sintchak MD, Rotstein SH, Murcko MA, Boger J, Thomson JA, Fitzgibbon MJ, Black JR, Navia MA. 1995. Comparative X-ray structures of the major binding protein for the immunosuppressant FK506 (tacrolimus) in unliganded form and in complex with FK506 and rapamycin. *Acta Crystallogr D 51*:511–521.
Yang D, Rosen MK, Schreiber SL. 1993. A composite FKBP12–FK506 surface that contacts calcineurin. *J Am Chem Soc 115*:819–820.

Note added in proof

A structure determination of the native FKBP12-FK506 complex bound to calcineurin has now been reported (Griffith JP, Kim JL, Kim EE, Sintchak MD, Thomson JA, Fitzgibbon MJ, Fleming MA, Caron PR, Hsiao K, Navia MA. 1995. X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex. *Cell 82*:507–522). A preliminary fit of mutant FKBP12 complex structures to the native FKBP12 in the calcineurin complex is entirely consistent with the results presented in this paper.

Appendix B

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 267, No. 4, Issue of February 5, pp. 2393-2399, 1992
Printed in U.S.A.

Tolerance of T4 Lysozyme to Proline Substitutions within the Long Interdomain α-Helix Illustrates the Adaptability of Proteins to Potentially Destabilizing Lesions*

(Received for publication, September 13, 1991)

Uwe H. Sauer‡, Sun Dao-pin§, and Brian W. Matthews¶

From the Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, University of Oregon, Eugene, Oregon 97403

To investigate the ability of a protein to accommodate potentially destabilizing amino acid substitutions, and also to investigate the steric requirements for catalysis, proline was substituted at different sites within the long α-helix that connects the amino-terminal and carboxyl-terminal domains of T4 lysozyme. Of the four substitutions attempted, three yielded folded, functional proteins. The catalytic activities of these three mutant proteins (Q69P, D72P, and A74P) were 60–90% that of wild-type. Their melting temperatures were 7–12 °C less than that of wild-type at pH 6.5. Mutant D72P formed crystals isomorphous with wild-type allowing the structure to be determined at high resolution. In the crystal structure of wild-type lysozyme the interdomain α-helix has an overall bend angle of 8.5°. In the mutant structure the introduction of the proline causes this bend angle to increase to 14° and also causes a corresponding rotation of 5.5° of carboxyl-terminal domain relative to the amino-terminal one. Except for the immediate location of the proline substitution there is very little change in the geometry of the interdomain α-helix. The results support the view that protein structures are adaptable and can compensate for potentially destabilizing amino acid substitutions. The results also suggest that the precise shape of the active site cleft of T4 lysozyme is not critical for catalysis.

Phage T4 lysozyme is a small monomeric protein with its structure divided into two distinct domains (Fig. 1). The active site is located at the junction of the two domains, and it might be expected that the alignment of one domain relative to the other would be critical for catalytic activity (*cf.* Storm and Koshland, 1970, but see also Jenks, 1969; Knowles, 1991). On the other hand, the crystal structure of a fully active mutant of T4 lysozyme has recently been described in which there is substantial variability in the "hinge-bending angle" between one domain and the other (Faber and Matthews, 1990).

To investigate the ability of T4 lysozyme to compensate for disruptive changes in its structure and to determine the need to conserve the alignment of the active site cleft, a series of proline substitutions was made in the long interdomain α-helix. Studies of proline substitutions and proline replacements in proteins and peptides include Matthews *et al.* (1987), Alber *et al.* (1988), O'Neil and DeGrado (1990), Strehlow *et al.* (1991), and Consler *et al.* (1991), among others.

The amino acids chosen for substitution with proline, Gln-69, Val-71, Asp-72 and Ala-74, are located in the middle of the long α-helix (residues 60–80) that connects the two domains of T4 lysozyme (Fig. 1). It was expected that the substitution of a proline at any of these sites would tend to significantly distort the α-helix and therefore change the alignment of the "upper" and "lower" domain. By making a series of replacements it was anticipated that the active-site cleft would be distorted in different ways. Also by including different substitutions it was possible to include sites that were both buried and solvent-exposed.

Gln-69 is largely exposed to solvent (Table I, Fig. 1) and its side chain does not obviously participate in stabilizing interactions with other parts of the protein. Val-71 is almost but not entirely buried. Its side chain makes many contacts with the non-polar residues Ile-3, Phe-4, Leu-7, Phe-67, Ala-74, Ile-100, and Phe-104 that contribute to the hydrophobic core both within the carboxyl-terminal domain of T4 lysozyme and connecting one domain with the other. Asp-72 is very solvent exposed and its side chain is located on the "backside" of the interdomain α-helix relative to the rest of the lysozyme molecule (Fig. 1). Ala-74 is largely, but not entirely buried, and makes non-polar contacts with residues Val-71, Ile-100, Val-103, and Phe-104 that contribute to the hydrophobic core in the carboxyl-terminal lobe of the molecule. In wild-type lysozyme Asp-70 makes an unusually strong salt bridge with His-31 (Anderson *et al.*, 1990). Asp-70 also accepts a hydrogen bond from the backbone amide of Ala-74. Since the replacement of Ala-74 with proline eliminates any possibility of hydrogen bonding to the amide it was likely that this replacement would perturb the Asp-70. . .His-31 salt bridge.

EXPERIMENTAL PROCEDURES

Mutagenesis—The proline mutations were introduced by site-directed mutagenesis according to the uracil template method developed by Kunkel (Kunkel, 1985; Kunkel *et al.*, 1987). The cysteine-free "pseudo wild-type" lysozyme (WT*),[1] in which the 2 cysteine residues present in wild-type had been replaced in order to facilitate thermodynamic measurements (Wetzel *et al.*, 1988; Matsumura and Matthews, 1989; Pjura *et al.*, 1990) was used as the reference protein. The lysozyme e-gene contained within a 630 base-pair *Bam*HI-*Hin*dIII fragment had been previously cloned into phage M13 mp18 yielding the derivative M13 mp18 T4e C54T C97A. It was then transformed into *Escherichia coli* strain CJ236 (dut, ung-, thi1, relA/pCJ105(CM'))

* This work was supported in part by National Institutes of Health Grant GM21967 and the Lucille P. Markey Charitable Trust. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstrasse 1, W-6900 Heidelberg, Germany.

§ Present address: Laboratory of Molecular Biology, NIDDKD, National Institutes of Health, Bldg. 2, Rm. 316, Bethesda, MD 20892.

¶ To whom correspondence should be addressed.

[1] The abbreviation used is: WT*, pseudo wild-type.

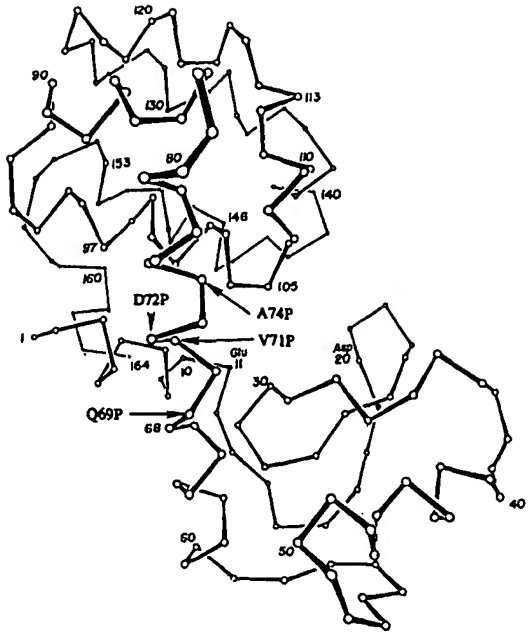


FIG. 1. **Backbone of T4 lysozyme showing the locations of the proline substitutions discussed in the text.** Mutations Q69P, D72P, and A74P give active, folded protein.

which was used to prepare uracil-containing single-stranded template DNA. After annealing of the mutagenic primer to WT* template DNA, circularization was accomplished by using the Klenow fragment of *E. coli* DNA polymerase I and T4 DNA ligase. The double-stranded DNA was subsequently used to transform competent *E. coli* JM101 cells. Sequencing of the whole lysozyme gene was carried out to confirm that no changes had occurred other than the ones introduced.

Protein Purification—Following subcloning into the plasmid expression vector pHN1403 (Muchmore *et al.*, 1989; Poteete *et al.*, 1991) 100 ml of cells grown overnight in 100 ml of LBH broth (10 g of tryptone, 5 g of NaCl, 1 ml of 1 M NaOH/liter) was added to 3 liters of LB broth (12 g of tryptone, 5 g of yeast extract, 10 g of NaCl, 1 g of glucose/liter) and grown at 37 °C under constant agitation (700 rpm) and air flow (12 liters/min) in a 5-liter fermenter until the optical density at a wavelength of 595 nm reached a value of 1.2. The temperature was then decreased to 26 °C, agitation and aeration were reduced to 200 rpm and 7 liters/min, respectively. Isopropyl-β-thiogalactoside (800 mg) was added to the growth media in order to induce lysozyme expression which was allowed to proceed for 100 min under continued stirring and aeration. The cells were then harvested into a 5-liter Erlenmeyer flask where they lysed. A few grains of DNase I were added to the now thick and viscous lysate which was left stirring at 4 °C for 2.5 h. The lysed cell suspension regained almost the viscosity of LB broth, was then taken out to room temperature, and placed on a magnetic stirrer for 30 min to allow for complete cell lysis. This step almost doubled the final yield of proline-containing mutant lysozymes. The lysate became more viscous again and was placed back at 4 °C, stirring for another 1.5 h until it regained the fluidity of LB broth. All the subsequent purification steps were carried out at 4 °C. After centrifugation at 10,000 rpm (17,700 × *g*) for 2 h, only the supernatant contained mutant lysozyme. It was dialyzed against 20 mM phosphate buffer at pH 6.5 until the conductivity reached a value of 3.8 mS/cm (1 Siemens = 1 Ω^{-1}). The dialyzed supernatant was loaded onto a 2.5 × 10-cm CM-Sepharose ion-exchange column which had previously been equilibrated with 400 ml of 50 mM Tris buffer at pH 7.3. Mutant lysozyme was eluted using an 800-ml salt gradient in the range from 50 to 300 mM NaCl in 50 mM Tris buffer, pH 7.3. Elution was monitored at a wavelength of λ = 280 nm. Peak fractions (absorbance at a wavelength of 280 nm above 0.4 units) were pooled, dialyzed against 50 mM phosphate buffer at pH 5.8 for 12 h, and concentrated on a 1 × 5-cm SP-Sephadex column previously equilibrated with the same buffer. The protein was eluted with SP buffer (100 mM $NaPO_4$, pH 6.5, 550 mM NaCl, 0.02% NaN_3) and stored in this buffer at 4 °C. Based on sodium dodecyl sulfate and high performance liquid chromatography analysis the purity of the mutant T4 lysozyme was estimated to be over 95%. Final amounts of 60–120 mg of mutant T4 lysozyme proteins were obtained. Wild-type T4 lysozyme typically yields about 150 mg/3-liter preparation.

The activity of the mutant lysozyme was measured at room temperature using the turbidity assay described by Tsugita (Tsugita *et al.*, 1968). Since absolute rates were not reproducible, activities were normalized to a wild-type control.

Measurement of Thermal Stability—Thermal denaturations at pH 2.0 and 6.5 were monitored by circular dichroism (CD) at a wavelength of λ = 229 nm on a Jasco J-500C spectropolarimeter as a function of temperature (Elwell and Schellman, 1975; Dao-pin *et al.*, 1990). The temperature was varied from 0 to 75 °C at a rate of 1 °C/min with a Hewlett-Packard 89100A temperature controller interfaced to a Hewlett-Packard 87 XM computer. The protein concentration was adjusted to 0.02 mg/ml by measuring the absorbance at λ = 280 nm using a double beam Varian 2290 spectrophotometer. A probe immersed in the sample solution just above the UV beam recorded the temperature. The solution was continuously stirred with a magnetic stirbar. Buffer solutions (150 mM KCl, 10 mM HCl at pH 2.0 and 150 mM KCl, 10 mM potassium phosphate at pH 6.5) were prepared from doubly deionized, degassed H_2O and were filtered before use through a 22-μm Millipore filter unit.

Thermal denaturations were repeated at least three times. Measurements of the mutants were flanked by WT* thermal denaturations under exactly the same conditions in order to minimize errors (Dao-pin *et al.*, 1990). The data were analyzed using standard van't Hoff techniques (Becktel and Schellman, 1987; Dao-pin *et al.*, 1990).

Crystallographic Methods—Crystal growth was attempted using both hanging-drop[2] as well as batch methods (Weaver and Matthews, 1987; Alber and Matthews, 1987) under conditions similar to those used for wild-type lysozyme. Both methods yielded crystals of D72P isomorphous with WT in the space group $P3_221$.

Refinement was carried out using the TNT package of refinement programs (Tronrud *et al.*, 1987). The positional coordinates and the temperature factors were refined simultaneously using the "conjugate directions" option in TNT which improves convergence.[3]

The starting model for refinement was the refined structure of the cysteine-free wild-type (WT*) (Pjura *et al.*, 1990; Bell *et al.*, 1991)[2] with residue 72 truncated to Ala. The general approach was to begin with low resolution (8–4 Å) rigid body refinement with the molecule considered as a single unit. This was followed by further rigid body refinement but with the mutant molecule divided into two parts (residues 1–80 and 81–162). Finally, the molecule was divided into three blocks (residues 1–59, 60–80, and 81–162). After examining the model on the graphics terminal (Jones, 1978), proline was built at position 72 and water molecules from the WT* structure were included in the model. Several cycles of positional refinement using moderately weighted geometric constraints were performed using data between 20 and 1.9 Å. Again, the model was inspected on the graphics terminal, some water molecules were added, others repositioned, and some side chains adjusted to better fit the electron density. Only those water molecules which had a final refined temperature factor of less than 80 $Å^2$, and, in addition, formed hydrogen bonds to protein or other bound water molecules and had no steric clashes were retained. Thereafter, several cycles of simultaneous positional and temperature factor refinement were alternated with model building until the crystallographic residual converged. The number of solvent molecules in the refined model is roughly the same as for WT* lysozyme, *i.e.* about 145. The refined coordinates have been deposited in the Brookhaven Data Bank.

RESULTS

Expression of Mutant Proteins

Gln-69 → Pro—Mutant Q69P could be expressed and purified in a straightforward manner, yielding up to 100 mg of protein from 4 liters of culture medium. The activity of the protein is close to that of wild-type (Table I), but it is less stable, with melting temperature 12.9 °C less than wild-type at pH 2.2 and 7.6 °C lower at pH 6.5. This behavior corresponds to that of a typical "temperature sensitive" mutant of

[2] A. E. Eriksson, W. A. Baase, and B. W. Matthews, manuscript in preparation.

[3] D. E. Tronrud, submitted for publication.

TABLE I

Sites of proline replacement

Amino acid	Fraction of side chain accessible to solvent	Mutant	Relative activity
			%
Gln-69	0.75	Q69P	88
Val-71	0.06	V71P	[a]
Asp-72	0.76	D72P	57
Ala-74	0.04	A74P	66

[a] Purified protein was not obtained for V71P.

TABLE II

Thermal stability of proline-containing mutants

T_m is the melting temperature and ΔT_m the difference between the melting temperature of the mutant and that of the pseudo wild-type lysozyme (see text). $\Delta\Delta G$, the difference between the free energy of unfolding of the mutant and pseudo wild-type lysozyme, was estimated from the relationship $\Delta\Delta G = \Delta S \cdot \Delta T_m$ (Becktel and Schellman, 1987) where ΔS is the entropy of unfolding of the wild-type protein (257 and 378 cal/degree mol at pH 2.0 and 6.5). This relationship may not be reliable when ΔT_m is large. The quoted values of $\Delta\Delta G$ are subject to error, estimated as ±0.5 kcal/mol for Q69P and D72P and ±1 kcal/mol for A74P. The estimated error in T_m is ±0.3 °C for Q69P and D72P and ±0.5 °C for A74P.

Protein	pH	T_m of mutant	T_m of WT*	ΔT_m	$\Delta\Delta G$
			°C		*kcal/mol*
Q69P	2.0	25.6	38.5	−12.9	−3.3
	6.5	55.6	63.2	−7.6	−2.9
D72P	2.0	28.3	38.5	−10.2	−2.6
	6.5	56.3	63.4	−7.1	−2.7
A74P	3.5	39.3	57.5	−18.2	−5.7
	4.0	46.9	62.2	−15.3	−5.0
	5.5	53.3	65.7	−12.4	−4.5
	6.5	50.8	62.9	−12.1	−4.6

T4 lysozyme selected by the random screen of Streisinger *et al.* (1961) (*cf.* Grütter *et al.*, 1979, 1987; Hawkes *et al.*, 1984). When stored at 4 °C at about 50 mg/ml, Q69P tended to form white opalescent aggregates which dissolved on warming to room temperature. This process was reversible and seemed to have no effect on stability or activity. Small crystals of the protein were obtained, apparently non-isomorphous with wild-type.

Val-71 → Pro—When V71P DNA was transformed into *E. coli* and the ability of a bacterial extract tested to form a halo on isopropyl-1-thio-β-D-galactopyranoside lysis indicator plates, no halo could be seen at 37 °C. After 24 h at 4 °C, a small halo was visible but we cannot rule out the possibility that this might be due to a small amount of WT* present as an impurity. Attempts to purify V71P by the method described above, or by using a French press to break open the bacterial cell walls, or by an alternative method described by Dao-pin *et al.* (1991b) were all unsuccessful. We presume that V71P is very unstable and/or is rapidly degraded by proteolysis.

Asp-72 → Pro—As with Q69P this protein could be readily purified by the standard procedure, yielding up to 130 mg of protein from a 4-liter culture. Its activity is about 60% that of wild-type (Table I) and it is less stable (Table II), again, roughly comparable to a typical temperature-sensitive mutant. Crystals isomorphous with wild-type could be grown at 4 °C using both batch and hanging-drop techniques. The latter method gave the largest crystals, 0.5 × 0.65 × 0.3 mm, from 2.0 M phosphate at pH 7.1. Some apparently non-isomorphous crystals were also obtained at 15 °C but have not been examined.

Ala-74 → Pro—As for Q69P and D72P, A74P behaved normally and yielded ~140 mg of protein/preparation from a 4-liter culture. The apparent activity is about two-thirds that of wild-type (Table 2). This mutant is less stable than both D72P and Q69P (Table II). Because of the low stability the protein tends to be partially unfolded at low pH, so the stability measurements under these conditions are less reliable than at higher pH. Some crystals of this protein were obtained from 50 mM phosphate, pH 7.9, 16% PEG 8000 and are non-isomorphous with wild-type.

Structure of D72P

The crystals of D72P appeared to be somewhat more sensitive to radiation than wild-type lysozyme and did not diffract as well. For this reason the exposure time per frame on a Xuong-Hamlin area detector system operating with graphite-monochromated CuK_α radiation from a Rigaku generator (40 kV, 100 mA) was increased to 60 s, compared to the usual 30 s for WT. Under these conditions the data set to 1.9 Å was 89% complete (Table III).

The map showing the difference in density between D72P and WT* lysozyme is shown in Fig. 2*a*. The positive density feature confirms the addition of the pyrrolidine ring. Also there is negative density at the site previously occupied by the carboxylate of Asp-72. Positive and negative densities also indicate the movement of the carbonyl oxygen of Asn-68 away from the helix axis (Fig. 2*b*). A negative feature next to the carbonyl group of Asp-70 indicates a movement of the oxygen toward the helix axis. The side chain of Asp-70 also moves slightly, as does His-31 (not shown) maintaining the strong (Anderson *et al.*, 1990) His-31...Asp-70 salt bridge. The movement of His-31 together with slight adjustments occurring throughout the lower lobe give the impression that the lower part of the long α-helix and the lower domain move as an essentially connected unit.

There is, however, a movement of the upper domain relative to the lower one. This shift is shown in Figs. 3 and 4. If backbone atoms in the amino-terminal domain of the mutant structure (residues 13–59) are superimposed on the corresponding atoms in WT* lysozyme (Fig. 4*a*), they have root-mean-square discrepancy of 0.16 Å, which is essentially experimental error. Similarly, the respective backbone structures within the carboxyl-terminal domains are also well conserved (root-mean-square discrepancy of 0.18 Å for residues 81–162, Fig. 4*b*). This shows that the structures within

TABLE III

Data collection and refinement statistics

Data for WT* taken from Eriksson *et al.*[2]

Protein	WT* (C54T/C97A)	D72P
Data collection statistics		
Mode of data acquisition	Film	Area detector (Xuong-Hamlin)
Cell dimensions		
a, b (Å)	60.9	60.8
c (Å)	96.8	98.6
Resolution (Å)	1.75	1.9
Unique reflections	14,562	15,147
Completeness of data (%)	65.0	89.4
R_{merge} (on intensities) (%)	4.6	5.5
Refinement statistics		
Resolution limits (Å)	6.0–1.75	20.0–1.9
RMS[a] deviation from ideal values		
Bond length (Å)	0.015	0.013
Bond angle (°)	2.1	2.0
Crystallographic residual (%)	14.8	16.7

[a] Root-mean-square.

The Journal of Biological Chemistry jbc

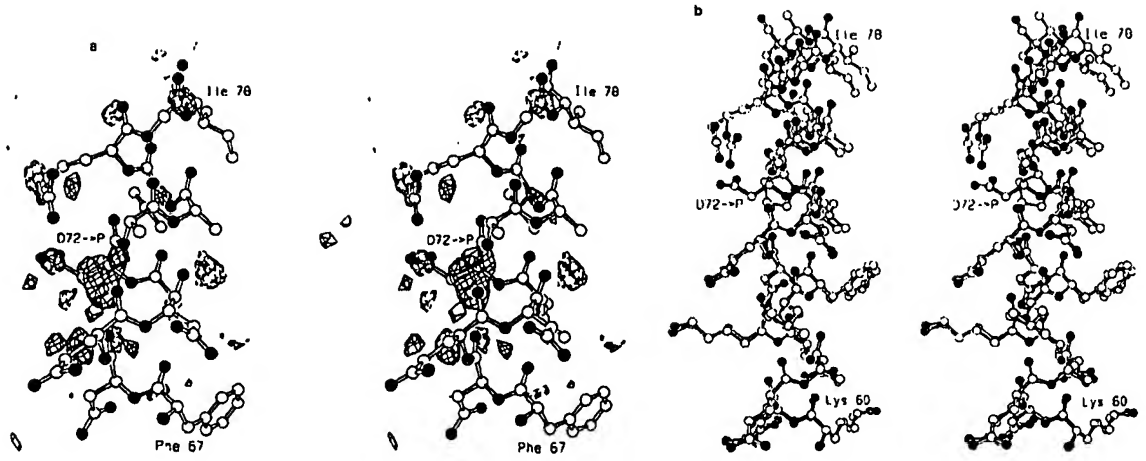


FIG. 2. *a*, map showing the difference in density between D72P lysozyme and wild-type. Amplitudes ($F_{Mut} - F_{WT}$) where F_{Mut} and F_{WT} are the structure amplitudes observed for the mutant and pseudo wild-type crystals. Phases calculated from the refined pseudo wild-type structure. Positive contours (*solid*) and negative contours (*broken*) drawn, respectively, at $+3\sigma$ and -3σ where σ is the root-mean-square value of the density throughout the unit cell. *b*, superposition of the A72P mutant structure (*solid bonds*) on wild-type (*open bonds*) in the vicinity of the mutation. The superposition of the coordinates shown were optimized by least squares prior to drawing the figure.

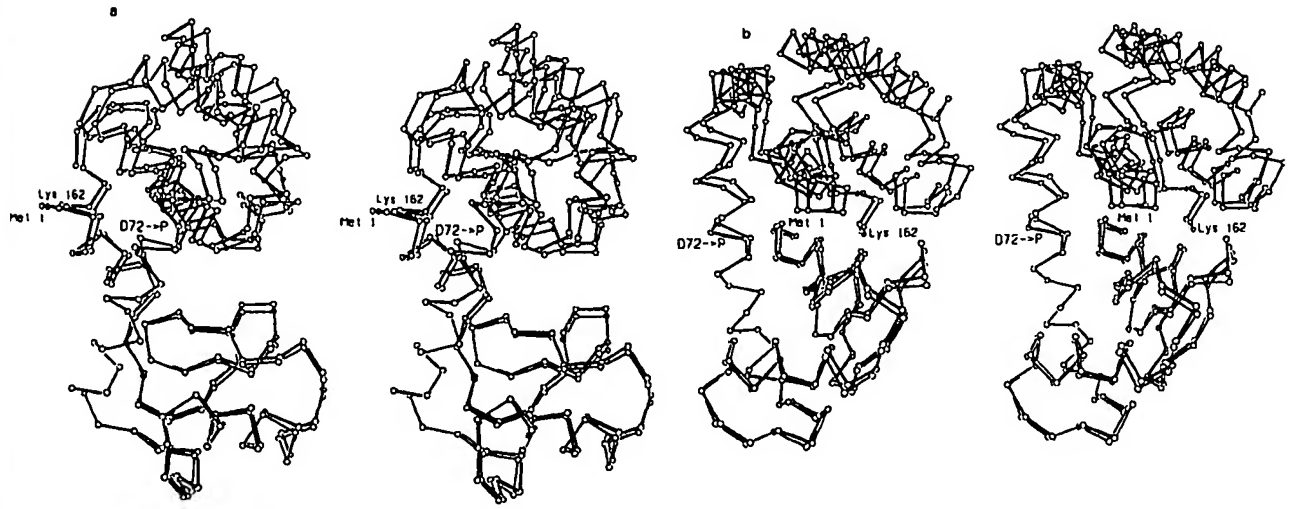


FIG. 3. *a*, backbone of the D72P mutant molecule (*dark bonds*) superimposed on WT* (*open bonds*). The figure is based on the optimal superposition of the amino-terminal region (residues 60–66) of the interdomain α-helix. *b*, superposition of D72P (*solid bonds*) on WT* (*open bonds*), as in Fig. 3*a* but rotated 90°.

the respective NH_2-terminal and COOH-terminal domains are conserved. It is the alignment of one domain relative to the other that changes in the mutant structure relative to WT*. As shown in Fig. 4, *a* and *b*, this movement corresponds to atom shifts up to about 1.5 Å. The pronounced maxima and minima in Fig. 4, *a* and *b*, are due to different distances of the corresponding atoms from the axis of rotation.

One way to analyze for movements of one part of the structure relative to another is to align the mutant structure with wild-type based on the superposition of relatively short segments of backbone. Two such alignments, based on 6-residue backbone segments on either side of the mutation site, are shown in Fig. 5, *a* and *b*. In Fig. 5*a*, in which the alignment is based on residues 60–66, the amino-terminal domain of the mutant structure agrees well with that of WT*. This shows that residues 60–66 and the NH_2-terminal domain are connected essentially as a rigid body. In contrast, when the superposition is based on residues 74–80 (Fig. 5*b*), neither the NH_2-terminal domains nor the COOH-terminal domains of the mutant and wild-type structures coincide. This suggests that the residues 74–80 are not connected rigidly to the COOH-terminal domain. The upper part of the 60–80 α-helix appears to be at least in part responsible for the observed flexibility. Small changes in this region can have large effects on the rest of the structure. The regions at the beginning and at the end of the long helix have been previously identified as "hinge-bending" regions in the M6I variant of T4 lysozyme (Faber and Matthews, 1990). The hinge-bending that was observed for the M6I mutant was assumed to be a low energy displacement because different molecules within the same crystal displayed different hinge-bending angles. In the case of M6I, the long interdomain helix appeared to remain rigid. In contrast, in the present case the long interdomain helix is bent relative to wild-type. Therefore the hinge-bending seen for D72P (Fig. 3) has a very different origin, and presumably very different energetics, relative to that seen for M6I.

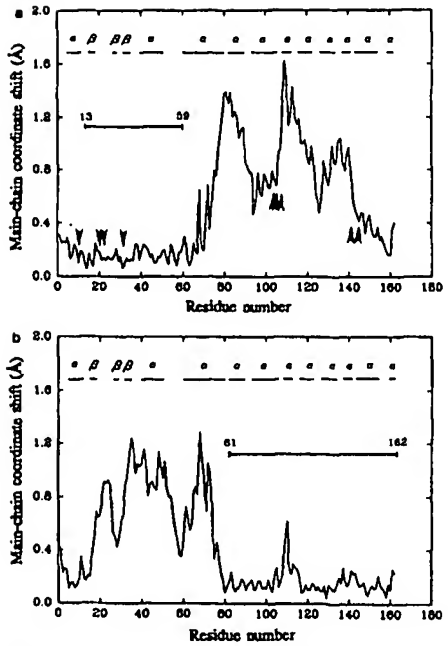


FIG. 4. **"Shift plots" showing the displacement between corresponding backbone atoms in the mutant D72P and wild-type lysozyme.** For each residue the value plotted is the root-mean-square discrepancy between the four backbone atoms in the mutant (N, C, CA, O) and the corresponding atoms in WT*. *a*, superposition of the structures of D72P and WT* based on residues 13–59 in the amino-terminal domain. The *arrowheads* indicate those residues that are presumed to contact an extended substrate (Glu-11, Asp-20, Thr-21, Leu-32, Phe-104, Gln-105, Thr-109, Thr-142 and Arg-145) (Anderson *et al.*, 1981). The contacts involving Thr-109 are to sugars in subsites A and B, whereas cleavage is between subsites D and E. *b*, superposition of D72P and WT* based on the carboxyl-terminal domain, residues 81–162.

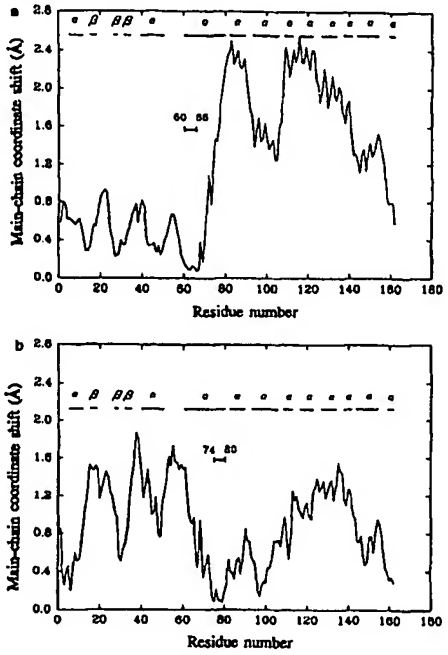


FIG. 5. *a*, superposition of D72P on WT* based on the alignment of the amino-terminal part of the interdomain α-helix, residues 60–66. *b*, superposition of D72P on WT* based on the alignment of the carboxyl-terminal part of the interdomain α-helix, residues 74–80. Note that *a* is similar to Fig. 4*a* indicating that residues 60–66 and the amino-terminal domain do not move very much relative to each other upon the proline replacement. *b*, however, is not very similar to Fig. 4*b*, showing that the carboxyl-terminal part of the interdomain α-helix does move somewhat relative to the carboxyl-terminal domain of the mutant protein.

Another way of detecting differences between two structures is by calculating a difference-distance matrix (Nishikawa *et al.*, 1972). All the intramolecular C_i^α-C_j^α distances are determined for one protein and then compared to the respective intramolecular distances of the second molecule. An example of such a plot is shown in Fig. 6. Since C_i^α-C_j^α distances in D72P were subtracted from the corresponding distances in WT*, positive contours (*solid lines*) reflect shorter distances in the mutant structure; negative contours (*dotted lines*) represent longer ones. As can be seen, the mutated residue 72 shows the largest shift with respect to residues 95, 122, and 157.

Fig. 2*b* illustrates the effect of the Asp-72 → Pro mutation on the long α-helix itself. In this figure residues 60–66 have been superimposed so as to make the structural change more obvious. The helix containing the proline has an overall bend of about 14°. A similar analysis of the structure of wild-type lysozyme, however, reveals that the same helix already has a bend of about 8.5°. Therefore the effect of the Asp-72 → Pro replacement is to increase the bending by about 5.5°.

Fig. 7 compares the hydrogen-bond distances within the interdomain helix in wild-type lysozyme and in the mutant structure. The first α-helical hydrogen bond is from the carbonyl oxygen of Thr-59 to the backbone amide of Ala-63. Not surprisingly, the introduction of the pyrrolidine ring at residue 72 results in a substantial increase in the distance between the nitrogen of residue 72 and the carbonyl oxygen of Asn-68. The nitrogen-oxygen distances for the successive residues are lengthened somewhat, but not excessively so (maximum 3.3 Å), suggesting that the hydrogen bonds within

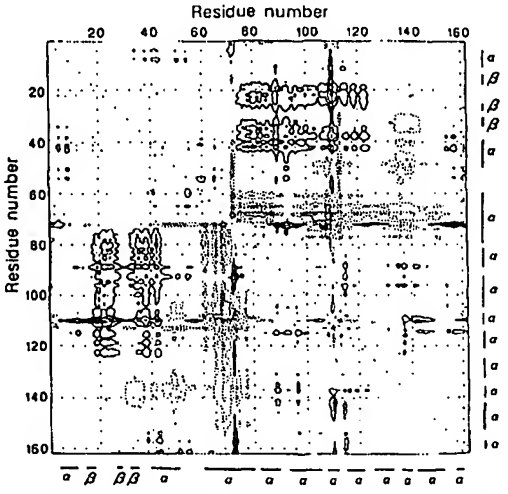


FIG. 6. **Difference distance matrix comparing distances between all pairs of α-carbon atoms in the mutant structure with the corresponding distances in WT*.** The quantity plotted is $\Delta_{ij} = r_{ij,WT*} - r_{ij,D72P}$ where $r_{ij,WT*}$ is the distance between the *i*th and *j*th α-carbon atoms in the structure of wild-type lysozyme, and $r_{ij,D72P}$ is the distance between the *i*th and *j*th α-carbon atoms in the D72P mutant structure. *Solid contours* are drawn at 0.3 Å, 0.6 Å, 0.9 Å, ... and indicate pairs of α-carbon atoms that are closer together in the mutant structure than in wild-type. *Broken contours*, drawn at −0.3 Å, −0.6 Å, −0.9 Å, ... indicate pairs of α-carbon atoms that are further apart in the mutant than in wild-type. Featureless regions indicate domains within which the structure in the mutant is essentially identical with that in wild-type.

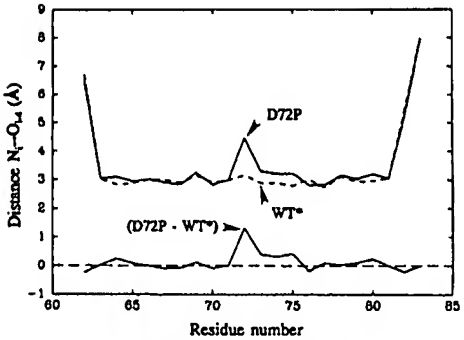


FIG. 7. **Hydrogen bond lengths within the interdomain helix.** The *upper plot* shows the distance between the nitrogen atom of residue *i* and the carbonyl oxygen of residue *i*-4. Within the α-helix (*i* = 63–81) this distance corresponds to a helical hydrogen bond. Values for the mutant, D72P, are indicated by the *solid line*; values for WT* are indicated by the *broken line*. The *lower plot* shows the difference in the distance between the mutant and wild-type structures.

TABLE IV

Ramachandran angles for residues within the long α-helix of T4 lysozyme

	D72P		WT*		Δφ	Δψ
	φ	ψ	φ	ψ		
Ile-58	−121.3	169.3	−126.2	167.2	4.8	2.1
Thr-59	−94.3	171.6	−90.9	166.4	−3.4	5.2
Lys-60	−62.9	−32.6	−56.8	−43.3	−6.1	10.7
Asp-61	−68.0	−35.9	−58.0	−45.7	−10.0	9.8
Glu-62	−69.1	−36.4	−62.3	−43.9	−6.8	7.5
Ala-63	−64.1	−41.9	−59.3	−46.1	−4.8	4.2
Glu-64	−70.4	−31.5	−69.5	−30.0	−1.0	−1.5
Lys-65	−61.1	−49.1	−69.1	−42.3	−8.1	−6.8
Leu-66	−61.3	−43.2	−59.4	−42.0	−2.0	−1.2
Phe-67	−55.5	−44.8	−61.6	−47.5	−6.1	−2.7
Asn-68	−62.4	−22.8	−56.2	−43.4	−6.2	20.6
Gln-69	−89.9	−38.9	−64.8	−40.1	−25.1	1.3
Asp-70	−66.6	−38.8	−67.8	−37.5	1.2	−1.4
Val-71	−63.8	−54.5	−68.1	−44.4	4.3	−10.1
Pro-72	−59.4	−34.8	−55.9	−48.8	−3.5	14.0
Ala-73	−64.1	−36.0	−60.7	−43.3	−3.4	7.3
Ala-74	−59.9	−50.2	−57.1	−56.2	−2.9	6.0
Val-75	−56.6	−48.0	−53.5	−49.7	−3.1	1.7
Arg-76	−58.0	−4.20	−64.4	−35.2	6.4	−6.8
Gly-77	−62.5	−36.3	−63.4	−45.1	0.9	8.9
Ile-78	−64.6	−46.0	−57.6	−46.3	−7.0	0.3
Leu-79	−65.1	−17.3	−69.5	−20.3	4.4	2.9
Arg-80	−97.3	−11.6	−95.9	−6.4	−1.4	−5.2
Asn-81	−93.5	119.8	−92.7	127.3	−0.9	−7.5

the remainder of the α-helix are maintained. In the mutant structure a solvent molecule is observed 3.9 Å from the carbonyl oxygen of Asn-68, suggesting a weak hydrogen-bonding interaction.

The (φ, ψ) angles of the residues within the long helix are listed in Table IV. The values observed for the proline (φ = −59.6°, ψ = −34.6°) agree very well with the average value (φ = −61°, ψ = −35°) for prolines in other protein structures (MacArthur and Thornton, 1991). At the site of the substitution, the addition of the pyrrolidine ring necessitates virtually no change in φ. The bigger change (Δψ = 14°) is in the successive peptide. Not surprisingly, the largest changes in (φ, ψ) are for the peptide between residues 68 and 69, which is in the previous turn of the α-helix and for which the hydrogen bond to the amide of residue 72 is disrupted.

DISCUSSION

The most striking result of the present study is the finding that proline residues can be substituted at several positions within the long interdomain α-helix of T4 lysozyme with only modest effects on catalytic activity. The proteins are destabilized relative to wild-type, but still fold and behave essentially normally. Attempts were made to substitute prolines at four sites and in three cases a functional protein was obtained. It is not as if there is one particular site at which a proline can be accepted. Rather, the data suggest that it may be possible to introduce prolines at additional sites within the interdomain helix, if not at many other sites in the protein as well.

The decrease in stability observed for the two mutants D72P and Q69P is very comparable with that found for temperature-sensitive mutants of T4 lysozyme such as R96H, T157I, and A98V identified by the random genetic screen of Streisinger *et al.*, 1961; Grütter *et al.*, 1969, 1987; Weaver *et al.*, 1989; Dao-pin *et al.*, 1991a). The mutant A74P is, however, less stable than any of these previously described variants. The decrease in stability associated with the proline substitutions seems to be associated to some extent with the inaccessibility of the residue to solvent but the correlation is not perfect. Val-71 is largely buried and a proline replacement at this site did not yield a functional protein. Ala-74 is also largely solvent inaccessible. In addition the proline substitution disrupts the α-helical hydrogen bond between residues 70 and 74 which, in turn, is likely to misalign and perhaps weaken the very strong salt bridge between Asp-70 and His-31 (Anderson *et al.*, 1990). In this case protein was obtained although with substantially reduced stability. O'Neil and DeGrado (1990) found that the energy cost of an alanine to proline replacement within a dimeric α-helical model peptide was 3.4 kcal/mol. Also Yun *et al.* (1991) found, by free energy simulations, exactly the same value for an alanine to proline replacement within a short polyalanine helix. These values are roughly comparable with those found here (average values of 3.2, 2.7, and 5.2 kcal/mol) but there is no reason to expect close agreement since the context is different in every case.

In the case of A72P, for which the crystal structure is available, the pyrrolidine ring is seen to introduce three unfavorable contacts with neighboring atoms (2.93 Å between C^δ and the peptide nitrogen of Val-71; 3.04 Å between C^δ and the peptide nitrogen of Val-71; 3.00 Å between C^δ and the carbonyl oxygen of Asn-68). Each of these contacts corresponds to an unfavorable van der Waals interaction energy in the range 1–2 kcal/mol (Levitt, 1974). These steric clashes are, presumably, a major factor in the destabilization of the mutant structure relative to wild-type.

The results provide further evidence that protein structures are adaptable and can compensate for amino acid substitutions at many sites (Matthews, 1987; Sondek and Shortle, 1990). It also illustrates the redundancy that is present in the amino acid sequence of a protein. Not every amino acid in the linear sequence is necessary for folding (Reidhaar-Olson and Sauer, 1988; Zhang *et al.*, 1991). Amino acid substitutions that are expected to distort and destabilize the folded structure of the protein, and disrupt a major α-helix that might be a key folding intermediate, do not prevent the formation of a folded functional protein.

Functional proteins were obtained with prolines substituted at positions 69, 72, and 74. Residues 69 and 72 are located in successive turns on the same side of the helix. Residue 74, however, is on the *opposite* side of the α-helix. Therefore one cannot argue that prolines are only accommodated on one side of the α-helix such that each substitution bends or distorts the α-helix in the same direction.

The substitution of a proline in an α-helix that is part of a protein is not the same as a proline substitution of an isolated

α-helix. In the absence of any constraints, a substituted proline is likely to completely disrupt the helix (Strehlow *et al.*, 1991) or to introduce a bend of up to 45° (Barlow and Thornton, 1988; Karle *et al.*, 1991). The structural consequences of a proline introduced into an α-helix in a protein, however, will be modulated by interactions between the α-helix and the rest of the protein. Prolines in α-helices in known protein structures are associated with bend angles that average 26 ± 5° (Barlow and Thornton, 1988), but this is not to say that an introduced proline will cause a change of this magnitude. The determination of the structure of D72P provides an example of the structural compromise that can result. The presence of the proline increases the bend angle of the α-helix, but only by 5.5°. Thus, it is not to be expected that each individual proline substitution will cause a major rearrangement of the two domains in T4 lysozyme. Nevertheless it can be anticipated that the different substitutions will cause distinct changes in the alignment of one domain relative to the other. This suggests that the precise shape of the active-site cleft in the resting enzyme is not critical for catalysis. At the same time it should be noted that the key catalytic residues are at the base of the active site cleft, and the structural changes in this region may therefore be relatively small (≤0.5 Å). The residues in T4 lysozyme that are presumed to contact an extended oligosaccharide substrate are indicated in Fig. 4*a*.

It should be noted that the crystals that were obtained for the replacement Asp-72 → Pro are isomorphous with wild-type. It is possible that the formation of such isomorphous crystals constrains the structure of D72P to be more similar to WT* than is the case in solution. The differences between the structures of D72P and WT* in solution may therefore be larger than those seen in Figs. 3 and 4. The mutants Q69P and A74P give crystal forms that are not isomorphous with wild-type. Hopefully, the structure analysis of these crystals will give a better overall impression of the structural changes that are induced by the proline substitutions.

The first determination of the structure of a temperature-sensitive mutant lysozyme (Grütter *et al.*, 1979) showed that relatively large changes in stability were accompanied by minimal changes in structure, these being localized to the immediate vicinity of the substituted amino acid. The ability of T4 lysozyme to accommodate destabilizing mutations in this manner has subsequently been seen on a number of occasions (Matthews, 1987) and is further exemplified by the present study.

Acknowledgments—We thank Drs. Walt Baase, Larry Weaver, Eric Anderson, and Elisabeth Eriksson for advice and help with, respectively, thermodynamic measurements, data collection, NMR measurements, and crystallographic refinement. Joan Wozniak and Sheila Pepiot also provided advice in protein purification, and D. W. Heinz and X-J. "Cai" Zhang helped with data analysis. We also thank Dr. J. A. Schellman for facilities.

REFERENCES

Alber, T., and Matthews, B. W. (1987) *Methods Enzymol.* **154**, 511–533
Alber, T., Bell, J. A., Dao-pin, S., Nicholson, H., Wozniak, J., Cook, S., and Matthews, B. W. (1988) *Science* **239**, 631–635
Anderson, D. E., Becktel, W. J., and Dahlquist, F. W. (1990) *Biochemistry* **29**, 2403–2408
Anderson, W. F., Grutter, M. G., Remington, S. J., Weaver, L. H., and Matthews, B. W. (1981) *J. Mol. Biol.* **147**, 523–543
Barlow, D. J., and Thornton, J. M. (1988) *J. Mol. Biol.* **201**, 601–619
Becktel, W. J., and Schellman, J. A. (1987) *Biopolymers* **26**, 1859–1877
Bell, J. A., Wilson, K. P., Zhang, X-J., Faber, H. R., Nicholson, H., and Matthews, B. W. (1991) *Proteins Struct. Funct. Genet.* **10**, 10–21
Consler, T. G., Tsolas, O., and Kaback, H. R. (1991) *Biochemistry* **30**, 1291–1298
Dao-pin, S., Baase, W. A., and Matthews, B. W. (1990) *Proteins Struct. Funct. Genet.* **7**, 198–204
Dao-pin, S., Alber, T., Baase, W. A., Wozniak, J. A., and Matthews, B. W. (1991a) *J. Mol. Biol.* **221**, 647–667
Dao-pin, S., Soderlind, E., Baase, W. A., Wozniak, J. A., Sauer, U., and Matthews, B. W. (1991b) *J. Mol. Biol.*, **221**, 873–887
Elwell, M., and Schellman, J. (1975) *Biochim. Biophys. Acta* **386**, 309–323
Faber, H. R., and Matthews, B. W. (1990) *Nature* **348**, 263–266
Grütter, M. G., Hawkes, R. B., and Matthews, B. W. (1979) *Nature* **277**, 667–669
Grütter, M. G., Gray, T. M., Weaver, L. H., Alber, T., Wilson, K., and Matthews, B. W. (1987) *J. Mol. Biol.* **197**, 315–329
Hawkes, R., Grütter, M. G., and Schellman, J. (1984) *J. Mol. Biol.* **175**, 195–212
Jenks, W. P. (1969) *Catalysis in Chemistry and Enzymology*, McGraw-Hill, New York
Jones, T. A. (1978) *J. Appl. Crystallogr.* **11**, 268–272
Karle, I. L., Flippen-Anderson, J. L., Agarwalla, S., and Balaram, P. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 5307–5311
Knowles, J. (1991) *Nature* **350**, 121–124
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 488–492
Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367–382
Levitt, M. (1974) *J. Mol. Biol.* **82**, 393–420
MacArthur, M. W., and Thornton, J. M. (1991) *J. Mol. Biol.* **218**, 397–412
Matsumura, M., and Matthews, B. W. (1989) *Science* **243**, 792–794
Matthews, B. W. (1987) *Biochemistry* **26**, 6885–6888
Matthews, B. W., Nicholson, H., and Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6663–6667
Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) *Methods Enzymol.* **177**, 44–73
Nishikawa, K., Ooi, T., Isogai, Y., and Saito, N. (1972) *J. Physiol. Soc. Jpn.* **32**, 1331–1337
O'Neil, K., and DeGrado, W. L. (1990) *Science* **250**, 646–651
Pjura, P., Matsumura, M., Wozniak, J., and Matthews, B. W. (1990) *Biochemistry* **29**, 2592–2598
Poteete, A. R., Dao-Pin, S., Nicholson, H., and Matthews, B. W. (1991) *Biochemistry* **30**, 1425–1432
Reidhaar-Olson, J. F., and Sauer, R. T. (1988) *Science* **241**, 53–57
Sondek, J., and Shortle, B. (1990) *Proteins Struct. Funct. Genet.* **7**, 299–305
Storm, D. R., and Koshland, D. E., Jr. (1970) *Proc. Natl. Acad. Sci. U. S. A.* **66**, 645
Strehlow, K. G., Robertson, A. D., and Baldwin, R. L. (1991) *Biochemistry* **30**, 5810–5814
Streisinger, G., Mukai, F., Dreyer, W. J., Miller, B., and Horiuchi, S. (1961) *Cold Spring Harbor Symp. Quant. Biol.* **26**, 25–30
Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr.* **A 43**, 489–503
Tsugita, A., Inouye, M., Terzaghi, E., and Streisinger, G. (1968) *J. Biol. Chem.* **243**, 391–397
Weaver, L. H., and Matthews, B. W. (1987) *J. Mol. Biol.* **193**, 189–199
Weaver, L. H., Gray, T. M., Grütter, M. G., Anderson, D. E., Wozniak, J. A., Dahlquist, F. W., and Matthews, B. W. (1989) *Biochemistry* **28**, 3793–3797
Wetzel, R., Perry, L. J., Baase, W. A., and Becktel, W. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 401–405
Yun, R. H., Anderson, A., and Hermans, J. (1991) *Proteins Struct. Funct. Genet.* **10**, 219–228
Zhang, X-J., Baase, W. A., and Matthews, B. W. (1991) *Biochemistry* **30**, 2012–2017